Available online at www.sciencedirect.com

ScienceDirect

Procedia Engineering 125 (2015) 18 – 24

Procedia Engineering

www.elsevier.com/locate/procedia

The 5th International Conference of Euro Asia Civil Engineering Forum (EACEF-5)

Returns to scale in buildings construction costs: Indonesian cases

Andreas Wibowo[a,b,*]

[a] RIHS Ministry of Public Works Indonesia, Jalan Panyawungan Cileunyi Wetan Kabupaten Bandung, 40393, Indonesia
[b] Graduate Program in Civil Engineering Parahyangan Catholic University, Jalan Merdeka 30 Bandung, 40117, Indonesia

Abstract

This paper examines returns to scale in building constructions in Indonesia based on large sample sizes of different project types (i.e., hotel or apartment, hospital, office, campus, and plant). The analysis demonstrates that costs tended to vary with sizes at a constant rate, as shown by cost capacity factors close to unity, with the exception of campus cases that supported decreasing returns to scale. This finding affirms those of previous studies that non-constant returns to scale in cost-size relationships appear to weakly exist for building constructions. At the very least, it also implies that a simple unit-cost approach remains a reliable method for early cost estimates. This paper also identifies some potential issues associated with constant returns-to-scale applications.

Peer-review under responsibility of organizing committee of The 5th International Conference of Euro Asia Civil Engineering Forum (EACEF-5)

Keywords: cost-capacity factor; constant returns to scale; regression analysis; building construction; Indonesia.

1. Introduction

The relationship between costs and sizes in any production systems is rationally modeled as monotonic increasing functions. This argument should also apply to construction industries: larger project sizes would technically require more efforts and resources to complete, which lead to higher costs to incur. While this fact is self evident, a more compelling issue that merits attention of cost engineers is how the rates of increase in construction costs compare with those in sizes or capacities – returns to scale. Cost-size functions may exhibit increasing, constant, decreasing returns to scale or a combination of different modes of returns to scale for different cost or size ranges.

In process industries (e.g., chemical, pharmaceutical, petroleum, water treatment industries), the use of increasing returns-to-scale assumption has been quite common for early cost estimates. Under this assumption, one would expect

* Corresponding author. *E-mail address:* andreaswibowo1@yahoo.de

1877-7058
Peer-review under responsibility of organizing committee of The 5th International Conference of Euro Asia Civil Engineering Forum (EACEF-5)
doi:10.1016/j.proeng.2015.11.004

the estimated costs of building new larger facilities to run at lower rates than the increase in capacities; thus, unit costs would tend to further decline over larger sizes. The application of the wellknown six-tenths rule, for instance, suggests that doubling the capacity would increase the cost by only about 50%. However, a study by Remer and Wong [1] for airport terminal constructions shows the contrary that doubling the size needs more than double the implementation cost, thereby representing decreasing returns to scale.

This paper investigates the returns-to-scale behavior for building construction projects in Indonesia. Over the years, abundant articles have been published focusing on parametric cost modeling for buildings but little effort, let alone those using Indonesia's building cases, was devoted to exploring the relationships between costs and sizes. Understanding these relationships is of particular useful for preparing preliminary cost estimates at early project stages during which the levels of project definition are not sufficient and no better information is available for more accurate estimations. From the academic standpoint, the objective of the present paper is threefolds – it attempts to enrich body of literature in this area, updates and extends similar previous studies [2,3] using a larger dataset, and fuels scholarly debates on the interpretation of research findings for future studies.

2. Power factor model

Despite the lowest expected accuracy level because of limited available information, early cost estimates are extremely important for sponsoring organizations and project teams in engineering and construction projects [4,5]. These estimates assist the clients for setting budgets, predicting tender prices, and managing the designs to meet budgets [6]. Under American Association of Cost Engineers Cost Estimate Classification System, early cost estimates for a level of project definition ranging from 0 to 2% can be designated as Class 5—the bottom level of the system—with expected accuracies spanning from as high as +40%/–20% to as low as +200%/–100%. The typical estimating methods for these estimates are stochastic or judgments with independent variables generally something other than a direct measure of the units of items being estimated [7].

One of the widely used stochastic methods is the power factor model, also known as scale of operation method [8]. This model assumes that the costs of similar facilities of different sizes vary with the size raised to some power [9]:

$$\frac{C_Y}{C_X} = \left(\frac{Q_Y}{Q_X}\right)^m \tag{1}$$

where C_Y = estimated cost of facility of capacity Q_Y and C_X = known cost of facility of capacity Q_X, m = cost-capacity factor, $m \geq 0$. To obtain more distinct expressions for relevant costs per unit capacity (i.e. unit costs), Eq. (1) can be algebraically rewritten as [2]:

$$\frac{C_Y}{Q_Y} = \frac{C_X}{Q_X}\left(\frac{Q_Y}{Q_X}\right)^{m-1} \tag{2}$$

Given $Q_Y > Q_X$, it is straightforward to see that a factor of m greater than unity corresponds to decreasing returns to scale, equal to unity as constant returns to scale (CRS), and lower than unity as increasing returns to scale. This factor can be empirically determined by regressing the log transformed data of costs against the log transformed data of sizes and will be the slope of the regression line. One would naturally expect a condition of $0 < m < 1$ to be present for building constructions to justify economies of scale because of the following possible reasons: advantages over volume discounts for purchasing some resource items in larger quantities, the policy of diminishing rates of indirect costs, more pooled and efficient resource allocations, higher productivity, more advanced construction methods for projects of larger sizes, etc.

3. Data acquisition and analysis

This research benefited from the rich construction cost database system of Research Institute for Human Settlements (RIHS) under Indonesia's Ministry of Public Works and Public Housing. The fields set up in the database system include location, size, area, function, budget source (public, private), hotel class (for hotels only), number of stories (including basements), and some information on technical specifications (e.g., construction materials, types of foundations, sources of electricity power, elevators). At the time of writing this paper, the updating of database (2014) is still on progress, so this research used the complete data of 2013 that contained about 1,050 project data points across the country with a wide variety of building types, ranging from airport terminal to laboratory buildings. However, not entire data points and fields were usable for this research and a substantial amount of data had to be removed, primarily due to blank entries of the required fields. It is worth noting here that the cost data stored in the system is disclosed contractual values, including profits and overheads and not the true costs of projects which remain unknown and confidential for public. Sizes or capacities of projects were proxied by gross floor areas (GFA) in this research. Cost data were grouped by the type of building and those with a very limited number of samples were omitted for subsequent analysis.

3.1. Data normalization

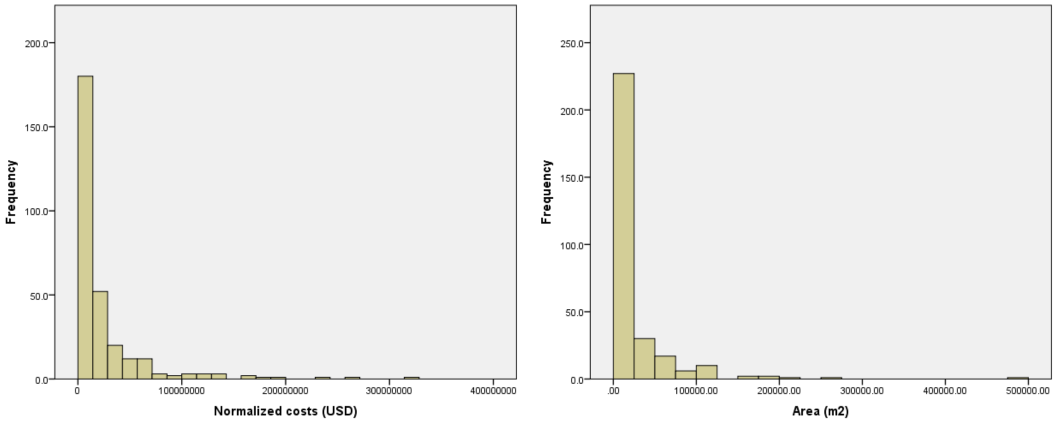
Because cost data may stem from different locations and years, they must be adjusted for location and time to ensure every observation is based on the same reference point. This adjustment process, also known as data normalization [10], involves a cost index, a dimensionless number which relates the cost of an item at a specific time (or location) to the corresponding cost at some arbitrarily specified time (or location) [9,11]. However, given that no detailed information on the start dates of projects were available in the system and the compiled data were from the same year, 2013, adjustments were only made for locations. To convert costs from one city to another city, the following formula was used:

$$C_j = C_i \frac{\text{cost index}_j}{\text{cost index}_i} \tag{3}$$

where C_j = cost at city j, C_i = cost at city i. Jakarta was selected as the reference city. Consumer price index (CPI) published by Statistics Indonesia (*Badan Pusat Statistik*, BPS) is not appropriate to use as cost index as it measures price changes of consumer goods and services not essentially relevant to construction activities and a misuse of this index could lead to misleading conclusions. Construction cost index (CCI; *Indeks Kemahalan Konstruksi*), also published by BPS, was deemed more suitable in the sense that it more reflects price changes in construction and was therefore used for this research; this index is a composite of 33 construction material prices, 6 heavy equipment rental costs, and 8 construction labor costs. Furthermore, unlike the BPS construction Wholesale Price Index (WPI; *Indeks Harga Perdagangan Besar*) that tracks price changes from year to year for the same location—a periodical index, CCI is of a spatial index that compares prices for different locations but same year.

3.2. Descriptive statistics of unit costs

Cost data normalized using Eq. (3) presented a considerably high dispersion in terms of costs and GFA. Provided that the existence of outliers can distort the regression results, it is thus essential to remove them from datasets. The first step to detect outliers was to calculate unit costs for each building type by dividing the normalized cost data points with their respective GFA. As had been anticipated, this process unveiled a great deal of nonsense data which were either extremely low or high. To secure more reliable data sets, this research applied a very restrictive criterion for every data point to be incorporated into regression analysis: it must fall within the range of the first quartile (Q_1) and the third quartile (Q_3) under the Tukey's hinges definition. This objective of reliability came at the expense of additional remarkable data omissions, leaving only useful 485 data points under 5 group categories: hotel/apartment, hospital, office, campus, and plant. Figs. 1a and b show, for instance, histograms of normalized costs and areas for hotel/apartment, respectively.



(a) (b)

Fig. 1. (a) Histogram of normalized costs for hotel/apartment; (b) Histogram of areas for hotel/apartment.

Table 1 demonstrates the key statistics of areas and normalized costs for these five building categories. As can be seen, the use of Q_1 and Q_3 as the lower and upper threshold, respectively, does not necessarily yield narrow data ranges. It comes as a no surprise if data distributions for both normalized costs and areas also present significant positive skewness coefficients (see also Figs. 1(a) and (b)), indicating that greater chances of extremely high costs and areas would be expected. This is because demand for larger scale construction projects tends to grow over time.

Table 2 shows the unit cost descriptive statistics of sampled projects.A closer examination on hospital cost data indicates that government funded hospital projects costed on average cheaper than privately funded projects (USD 614 per m^2, n = 28 versus USD 667 per m^2, n = 22). Nonethelesss, a comparable, meaningful analysis could not be exercised for other groups because of highly unequal sample sizes.

3.3. Cost capacity factors

Figs. 2(a) and (b) present regression straight lines and equations for hotel and plant building categories for examples. Table 3 depicts the key regression outputs for five building types. The analysis of variance advocates that the entire regression models were found to be statistically significant for at least at 0.0005 level (not shown). The remarkably high coefficients of determination (R^2)—close to 1 with the exception of campus category ($R^2 = 0.641$)––also suggest that these models have strong and very strong predictive capabilities. Both regression unstandardized coefficients (b) and constants were also significant for, at least, 0.01 level. These b coefficients are the respective cost capacity factors.

3.4. Discussion

The resulting m coefficients are somewhat surprising and unexpected as they suggested strong evidences of CRS ($m \approx 1$), denoting that costs tend to vary with sizes at a constant rate; only campus buildings were excepted for this conclusion as they exhibited decreasing returns to scale ($m > 1$). One may argue that the strict requirement for data inclusion adopted in this research improves the likelihood of CRS and relaxing the condition would likely change the type of returns to scale, either decreasing or increasing, depending on data distribution. While this argument might be plausible although it must be tested on datasets in any case, the research findings presented here are, nonetheless, to some degree consistent with those of previous studies. Wibowo and Wuryanti [2] and Amelia and Abduh [3] indicate that non-CRS appears to be weakly present for building constructions with the only exception being office buildings that remained elusive as these two studies arrived at discordant conclusions (see Table 4).

Table 1. Descriptive statistics of areas and normalized costs (2013) for sampled projects.

Statistics	Type of Building					
	Hotel/apartment (n = 297)	Hospital (n = 50)	Office (n = 82)	Campus (n = 11)	Plant (n = 45)	Total (n = 485)
Area						
Min. (m2)	3,000	420	442	700	800	420
Max. (m2)	480,000	40,000	121,000	6,000	200,000	480,000
Median (m2)	10,000	3,620	1,650	3,000	8,000	8,000
Mean (m2)	24,960	6,315	6,598	2,964	16,180	18,620
Std. Dev (m2)	42,748	7,659	15,549	2,186	30,055	36,292
Skewness	5.680	2.635	5.569	0.340	5.460	6.396
Cost						
Min. (USD)	1,851,271	236,141	184,399	243,553	341,359	184,399
Max. (USD)	327,868,850	21,711,115	49,000,000	2,167,729	91,047,041	327,868,850
Median (USD)	11,200,000	2,097,354	712,234	917,458	3,508,772	7,317,073
Mean (USD)	25,372,579	4,098,835	2,770,385	971,267	7,368,879	17,134,129
Std. Dev. (USD)	39,583,253	4,979,058	6,406,391	744,835	13,701,481	33,070,313
Skewness	3.942	2.301	5.376	0.547	5.436	4.781

Table 2. Descriptive statistics of unit costs (2013) for sampled projects

Type of Building	Min. (USD/m2)	Max. (USD/m2)	Mean (USD/m2)	Std. Dev. (USD/m2)
Hotel/apartment	597.184	1275.209	1043.465	162.410
Hospital	362.923	798.468	636.925	100.579
Office	361.288	509.400	422.600	32.710
Campus	304.205	361.288	326.583	22.096
Plant	423.313	515.274	456.305	30.995

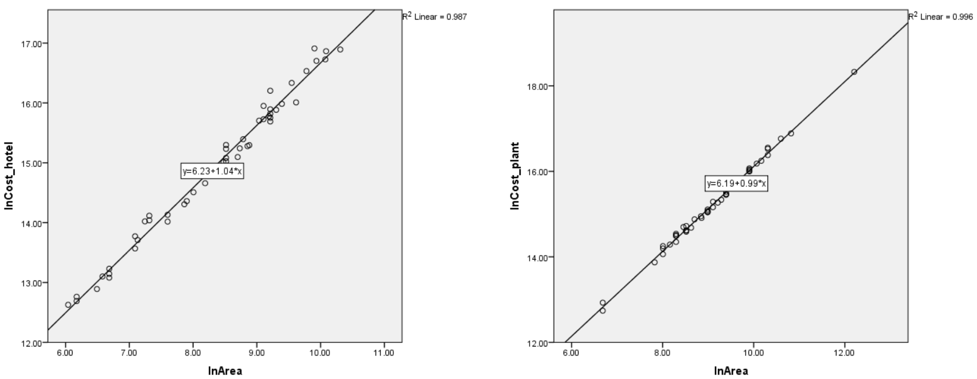


Fig. 2. (a) Regression for log-cost and log-area for hotel/apartment buildings; (b) Regression for log-cost and log-area for plant buildings.

Table 3. Key regression outputs.

Type of Building	R2	Unstandardized Coefficient (b)	Constant
Hotel/apartment	0.973	0.992a	7.015a
Hospital	0.979	1.012a	6.349a
Office	0.996	1.000a	6.045a
Campus	0.641	1.127a	4.642b
Plant	0.996	0.993a	6.188a

Note: a) Significant at least at 0.0005 level; b) Significant at least at 0.001 level

Table 4. Comparison of estimated cost capacity factors of previous studies.

Type of Building	Current Study	Remer and Wong [1]	Wibowo and Wuryanti [2]	Amelia and Abduh [3]
Hotel/apartment	0.992	NA	NA	NA
Hospital	1.012	NA	0.967	NA
Office	1.000	NA	0.836	1.360
Educational building (not specific)	NA	NA	0.984	NA
Elementary school building	NA	NA	NA	0.891
Campus (university)	1.127	NA	NA	1.004
Plant	0.993	NA	NA	NA
Airport terminal	NA	1.2a	NA	NA
	NA	2.9b	NA	NA

Note: a) terminal areas; b) passengers per year; NA = not available.

These research findings are also, to a large extent, in agreement with Latief, Wibowo, and Iswara [12] that attempted to develop a preliminary cost estimation model for high-rise residential buildings of Ex-Ministry of Public Housing. Running on 55 project data, their stepwise multiple linear regression model found 4, out of 22 initially identified, independent variables i.e. GFA, area per unit (APU), height of building (HOB), and type of foundation (TOF) to be the statistically significant predictors of project costs. Their model also satisfactorily fitted the data with R^2 of 0.953 and, interestingly, GFA had the supreme explanatory power with a standardized regression coefficient (β) of 0.964 whereas APU, HOB, and TOF had β coefficients of only 0.118, 0.098, and 0.088, respectively. To put it simply, GFA alone has been proven to be sufficient for explaining costs of high-rise residential building projects. Under the assumption of a linear relationship between costs and GFA, multiplying the latter with a constant cost rate results in a nearly accurate cost estimate, thereby supporting the existence of CRS.

Because research efforts on parametric cost modeling typically relied on contractual values as dependent variables, the application of CRS could lead to another question of interest with respect to the quality of bid proposals, however. Building contractors seem to be comfortable with average costs per m2 approach for estimating their final bid prices, do a fit and match for unit price analyse to justify their prices, and prepare more appropriate cost estimates only when successfully winning the contracts. It may also happen that indirect costs (i.e. mark-ups and overheads), which in many cases critically define the bidders' competitiveness, merely serve as a "plug" to close gaps between the desired final prices and direct costs. On a broader context, if this speculative presumption turns out to be correct, thus issues such as productivity and efficiency might be of less relevance in Indonesia's building construction industry which is definitely unfavorable under any circumstance for the sake of the industry itself. Moreover, this industry should discourage the practice of "get the project first and negotiate later."

4. Concluding remarks

This paper reports the existence of CRS in building constructions. This cost-size relationship would be very much helpful for preparing cost estimates in particular when project developments are still on early phases and the percentages of engineering complete are very low. However, this paper is in no way intended to be final. Its research findings are completely open for discussions and challenges, especially when larger sample sizes of high quality data are available. This paper also leaves avenues for future research with respect to factors that give rise to the tendency of building contractors to use a simple unit price approach for calculating their bid prices.

Acknowledgements

The author wishes to thank RIHS for allowing him to use the data from its database system.

References

[1] D.S. Remer, C. Wong, Cost scale-up factors for airport construction, Cost Eng. 38 (1996) 24–26.

[2] A. Wibowo, W. Wuryanti, Capacity factor based cost models for buildings of various functions, Civil Eng. Dimension 9 (2007) 70–59.

[3] A.F. Amelia, M. Abduh, Model faktor kapasitas biaya untuk estimasi biaya konseptual konstruksi bangunan gedung, Proc. of Int. Civil Eng. Conf., 25–26 August 2006, Petra Christian Univ. Surabaya, pp. 273–279.

[4] N. Fragkakis, S. Lambropoulos, G. Tsiambaos, Parametric model for conceptual cost estimation of concrete bridge foundation, J. Infr. System 17 (2011) 66–74.

[5] G.D. Oberlender, S.M. Trost, Predicting accuracy of early cost estimates based on estimate quality, J. Constr. Eng. Mgmt. 127 (2001) 173–180.

[6] D.J. Lowe, M.W. Emsley, A. Harding, Predicting construction cost using multiple regresssion techniques, J. Constr. Eng. Mgmt. 132 (2006) 750–758.

[7] American Association of Cost Engineers (AACE), Cost estimate classification system, AACE Int. Recommended Practice No. 17R-97 (1997).

[8] L.R. Dysert, T.W.Pickett, So you think you're an estimator? Cost Eng. 47 (2005) 30–35.

[9] H.C. Nelson et al., Capital investment cost estimation, in: K.K. Humphreys (Ed.), Jelen's Cost and Optimization Engineering, McGraw-Hill, New York, 1991, pp. 361–433.

[10] K. Phaobunjong, C.M. Popescu, Parametric cost estimating model for buildings, AACE Int. Transaction (2003), pp. EST.13.1–EST.13.11.

[11] American Association of Cost Engineers, Cost engineering terminology, AACE Int. Recommended Practice No. 10S-90 (2004).

[12] Y. Latief, A. Wibowo, W. Iswara, Preliminary cost estimation using regression analysis incorporated with adaptive neuro fuzzy inference system, Int. J. of Tech. 1 (2013) 63–72.